J Psycholinguist Res (2009) 38:75–92
DOI 10.1007/s10936-008-9084-0

Distributional Stress Regularity: A Corpus Study

David Temperley

© Springer Science+Business Media, LLC 2008

Abstract The regularity of stress patterns in a language depends on *distributional stress regularity*, which arises from the pattern of stressed and unstressed syllables, and *durational stress regularity*, which arises from the timing of syllables. Here we focus on distributional regularity, which depends on three factors. *Lexical stress patterning* refers to normal stress patterns within words; *interlexical stress patterning* refers to patterns that arise from word combinations; and *contextual stress patterning* refers to adjustments in normal lexical stress patterns (such as the well-known phenomenon of "stress clash avoidance"). A corpus study was done to assess the effect of these three factors on distributional stress regularity in conversational and formal spoken English, by comparing the degree of stress regularity in stress-annotated corpus data to randomly manipulated versions of the data and to "citation-form" stress patterns drawn from a phonetic dictionary. The results show that both lexical and interlexical patterning contribute significantly to stress regularity in English; contextual stress patterning does not, and in fact significantly reduces regularity in comparison to citation-form stress patterns.

Keywords Stress · Prosody · Corpus analysis · Syntactic choice

Introduction

Stress has been widely studied from both experimental and theoretical perspectives, and has proven to be a subtle and complex linguistic phenomenon (Fry 1958; Morton and Jassem 1965; Chomsky and Halle 1968; Liberman and Prince 1977; Selkirk 1984; Hayes 1995; Sluijter and van Heuven 1996). A recurrent issue in the study of stress has been the question

D. Temperley (✉)
Eastman School of Music, University of Rochester, 26 Gibbs St.,
Rochester, NY 14604, USA
e-mail: dtemperley@esm.rochester.edu

of regularity.[1] Traditionally, a distinction has been made between "stress-timed" languages, in which stresses are equally spaced in time, and "syllable-timed" languages, in which syllables are equal in duration, resulting in irregular timing of stresses (Pike 1945; Abercrombie 1967). Experimental research has not borne out this distinction. Studies by Roach (1982) and Dauer (1983) examined inter-stress intervals in several languages, including both supposedly stress-timed languages like English and supposedly syllable-timed languages like Spanish. In all languages studied, intervals between stresses varied quite widely, and increased with the number of unstressed syllables between stresses. However, Dauer did find that unstressed syllables tended to be shorter in intervals containing more of them, indicating some tendency toward regular stress timing. Other studies have found similar tendencies towards stress regularity; for example, in polysyllabic words with initial stress, syllables tend to be pronounced more quickly than in monosyllabic words (Lehiste 1972; Beckman et al. 1990). Most recent treatments of stress, while not endorsing the distinction between stress-timed and syllable-timed languages, suggest that there is some tendency toward regularity of inter-stress intervals, at least in English (Selkirk 1984; Ladefoged 1993; Hayes 1995; Cruttenden 1997).

Empirical studies of stress regularity have generally focused on the timing of syllables, and in particular on variations in syllable duration that enhance regularity; we will call this *durational stress regularity*. However, another factor also affects stress regularity, namely the actual pattern or distribution of stressed and unstressed syllables—what we will call *distributional stress regularity*. It can be seen that this could greatly affect the overall regularity of stresses. As two (hypothetical) extreme examples, consider a language in which unstressed (U) and stressed (S) syllables occurred in perfect alternation (1a), versus a language in which they were randomly arranged (1b):

(1a) S U S U S U S U S U S U S U S
(1b) S S U U U S U S S U U U S S S

Of course, the actual degree of stress regularity in a language also depends on the timing of the syllables. But under most reasonable assumptions about syllable timing, it would seem that (1a) would yield more regular stresses than (1b) (unless the syllables in (1b) were systematically timed to equalize the inter-stress intervals—and the abovementioned studies suggest that this does not occur).

What linguistic phenomena might contribute to distributional stress regularity? Three separate kinds of phenomena suggest themselves, all of which have been observed in English (and other languages):

1. Stress patterns within words, and the rules governing these, may favor a regular distribution of stresses. It is well-known that certain morphological patterns in English promote stress alternation. This is seen most clearly in cases of stress shift within stems as affixes are added, as in *inféstínfestátion*. In many theories of English word stress, the preference for stress alternation is explicit, reflected in rules which enforce or encourage alternating stress—for example, by adding stress to a syllable before an unstressed one or vice versa

[1] The terms *isochrony*, *eurhythmy*, and *euphony* are also sometimes used to describe stress regularity, but none of these are very satisfactory. *Isochrony*, strictly speaking, means exact regularity of events in time—as is found, for example, in some artificial experimental stimuli; obviously this rarely occurs in natural speech. *Euphony* is too general, as it refers to any kind of sonically pleasing linguistic usage. As for *eurhythmy*: Ironically, this word seems prosodically very awkward if not ill-formed. Trisyllabic nouns with the suffix $-y$ invariably have initial stress: *lunacy*, *eulogy*, etc. But the stem *rhyth*(*m*) is invariably stressed; to use it without stress feels quite wrong. (The prefix *eu-* is usually unstressed, though *eulogy* and *euphony* are exceptions.) Thus there seems to be no good pronunciation of *eurhythmy*.

(Chomsky and Halle 1968; Liberman and Prince 1977; Selkirk 1984; Hayes 1984, 1995). It appears, then, that lexical stress patterns within polysyllabic words might greatly affect the overall stress regularity of English. This factor in distributional stress regularity will be called *lexical stress patterning*.

2. Words may be chosen and combined in ways which enhance stress regularity. For example, there might be a tendency to avoid "stress clashes" between adjacent words, where a word with final stress is followed by one with initial stress. Factors favoring stress regularity might include grammatical constraints on word order, idiomatic phrases, or spontaneous choices in sentence production. Some examples of word patterns influencing stress regularity are provided in an essay by Bolinger (1965). Bolinger suggests that many peculiar idiomatic expressions have arisen to avoid stress clash:

   (2a) The sailor is drunk/a drunken sailor (*not* a drunk sailor)
   (2b) half an hour (*not* a half hour)
   (2c) without a doubt (*not* without doubt)

   The phrase *drunk sailor* would cause a stress clash; this might explain why the otherwise obsolete form *drunken* has been retained. Another example of how word patterns might favor stress regularity is offered by Kelly and Bock (1988). Kelly and Bock note that in English, two-syllable nouns are generally trochaic (strong-weak) while verbs are generally iambic (weak-strong). (This pattern is nicely illustrated by noun/verb homographs such as *pérmit*/*permít*.) In a corpus analysis, they show that nouns are significantly more likely to be followed by stressed syllables than verbs are, while verbs are more likely to be preceded by stressed syllables. It can be seen that these patterns favor a regular alternation of stressed and unstressed syllables. If we consider the typical case of a trochaic noun and an iambic verb, the most common stress patterns are also the most regular:

   (3a) nouns: S | S U (uncommon) S U | S (common)
   (3b) verbs: S | U S (common) U S | S (uncommon)

   The role of word patterns in distributional stress regularity will be called *interlexical stress patterning*.

3. It is generally assumed that every English word has a normal stress pattern which reflects its most common prosodic usage; lexical and interlexical stress patterning refer to patterns of stress that emerge when these normal prosodic forms are used. However, the normal stress patterns of words are sometimes adjusted to enhance stress regularity, and in particular to avoid stress clashes. For example, *thirtéen* becomes *thírteen mén*; *Chinése* becomes *Chínese róom*. Such phenomena have been quite extensively studied in English (Liberman and Prince 1977; Hayes 1984, 1995). Another situation where normal lexical stress patterns may be adjusted is where several unstressed syllables occur in a row. Consider the two sentences below:

   (4a) I gave it to the teacher
   (4b) I'm giving it to John

   Each of these sentences contains three successive syllables that would normally be unstressed (*it to the* in the first case, *-ing it to* in the second). But in the first case, there seems to be a slight tendency to stress *to*; in the second case, *it* feels slightly stressed. It can be seen that these adjustments produce a perfectly alternating stressed pattern in each case. It is possible that deviations from normal lexical stress patterns play a significant role in stress regularity. This factor in distributional stress regularity will be called *contextual stress patterning*.

The three kinds of patterning proposed above—lexical, interlexical, and contextual—are all logically independent; in principle, distributional stress regularity in a language might be enhanced by all or none of them or any combination. As noted above, all have been claimed to operate to some degree in English. However, there has been very little quantitative study of their actual effect on stress regularity. With regard to interlexical patterning, the corpus study of Kelly and Bock—showing that nouns and verbs occur in different stress environments—has already been mentioned; it is unclear how much this specific pattern contributes to overall stress regularity. Contextual stress timing has been examined in a study by Cooper and Eady (1986). In this study, participants read phrases such as *thirteen black boards* and *thirteen blackboards*. The "stress clash avoidance" view of Hayes (1995) and others predicts that *thirteen* would have final stress in the first case and initial stress in the second; no such effect was found. Beckman et al. (1990) did a similar study and similarly found no effect of stress clash avoidance. However, these studies used only a small number of phrases, and focused on one very specific type of contextual stress patterning. It remains to be seen whether contextual stress patterning in general plays any significant role in enhancing stress regularity.

In this study, we investigate the effect of lexical, interlexical, and contextual patterning on distributional stress regularity in English. We use corpora of natural speech annotated with stress information, manipulating the data to isolate the effects of the various factors. We begin by randomly rearranging the stressed and unstressed syllables of each sentence. Comparing this random pattern to the actual stress pattern of the annotated text, we can observe the overall stress regularity of the original data—the effect of lexical, interlexical, and contextual patterning in combination. We then create modified versions of the corpus in which the three kinds of stress patterning are systematically controlled. By replacing the annotated stresses with "citation-form" stress patterns—drawn from a phonetic dictionary—and randomly reordering the words of each sentence, we can examine the degree of stress regularity that occurs due to lexical patterning alone. By then restoring the original word order, but retaining the citation-form stresses for each word, we can examine the effect of interlexical patterning. And by comparing this to the original, annotated stress pattern, we can observe the effect of contextual patterning. In effect, these tests will tell us how much of the total regularity of the final output is due to lexical, interlexical, and contextual patterning.

Corpora

While many natural language corpora are now available, including a number of phonetically annotated corpora, very little corpus data is available with manually annotated stress information. Only two English corpora of this kind could be found. One is the Switchboard Transcription Project (henceforth the STP corpus). In this project, phonetic annotations were done on a subset of the Switchboard corpus, a corpus of several hundred informal telephone conversations (Greenberg 1997). At a later stage, a further subset of this phonetically annotated data—about 45 min—was annotated with stress information. In the phonetic annotations, syllables are explicitly marked; the stress annotations consist of the symbols 1 (for "fully accented"), 0.5 (for "intermediate accent"), and 0 (for "completely unaccented") appended to the vowel of each syllable (Hitchcock and Greenberg 2001).[2] (The data also contains a small number of syllabic consonants.) Two transcribers annotated the same data. The stress values for each transcriber are available as well as the averages between them; the average values are used here. (As these values represent averages of scores of 0, 0.5, or 1, they also include some values of 0.25 and 0.75.) The transcribers showed a high level of agreement; in distinguishing between unstressed syllables (0) and intermediately or fully stressed syllables (0.5 or 1), they agreed on 95% of syllables (Hitchcock and Greenberg 2001).

Spontaneous conversational English—such as that represented in the STP corpus—might be considered the most appropriate sort of data for our purposes, as it exemplifies the most "natural" and common kind of language production. However, it also has disadvantages. Spontaneous conversation is full of false starts, repairs, repetitions, and interruptions, which may distort the speaker's intended utterances and any stress regularity associated with them. For this reason, it seemed desirable to also include a corpus of more formal "correct" English. For this purpose we used the Boston University Radio News Corpus (henceforth the BU corpus), which includes phonetic transcriptions of several hours of radio news stories. A small subset of the corpus was hand-annotated with stress symbols (Ostendorf et al. 1995). In this corpus, only two levels of stress were used: stressed syllables are marked with 1, and unstressed syllables are unmarked. The corpus does not explicitly segment the text into syllables, so for present purposes we simply define syllables by vowels: Every unmarked vowel is taken to represent an unstressed syllable.

The current project also required a phonetic dictionary, showing "citation-form" pronunciations for each word. For this purpose we used the CMU Pronunciation Dictionary, version 6.[3] This dictionary contains about 127000 words, including a large number of proper names and separate entries for word forms such as verb forms and singular/plural noun forms. Syllables are marked with primary stress (1), secondary stress (2), or no stress (0). In many cases, alternative pronunciations are given for a word; in this case we use only the first pronunciation, which generally seems to be the most common.

Recall that the purpose of using a phonetic dictionary was to examine the amount of stress regularity that results simply from using the normal pronunciations of words, excluding stress-clash avoidance and other contextual effects. For the most part it seems reasonable to take dictionary pronunciations as representing normal word pronunciation. One important exception, however, concerns function words—words such as determiners, pronouns, prepositions, conjunctions, and auxiliaries. In the CMU dictionary, most function words are represented as stressed, as they would be if pronounced in isolation. In practice, however, function words are unstressed in a large majority of cases, though they may occasionally be stressed—for example, if a pronoun is used in narrow focus (*HE doesn't care*) or (sometimes) in phrase-final position (Selkirk 1996). That function words are normally unstressed seems to be the usual assumption in stress research (see for example Hayes 1995, p. 24), and is also confirmed by inspection of the data used here. Thus, a special version of the CMU dictionary was created in which function words were unstressed. The categorization of words into function words and content words is a non-trivial problem; this is discussed in the Appendix, and the list of words defined as function words is shown there. It should be noted that the procedure used here for finding citation-form stress patterns is imperfect, as it takes no account of syntactic categories, which sometimes affect stress patterns. However, inspection of the results suggests that the current procedure succeeded in identifying the

[2] Actually unstressed syllables were simply left unmarked, but it is clear from Hitchcock and Greenberg (2001) that they were regarded as having stress of 0 for the purpose of averaging the scores.

[3] Publicly available at www.speech.cs.cmu.edu/cgi-bin/cmudict.

correct citation-form stress patterns for the vast majority of word tokens (see the Appendix for discussion).

The stress-annotated STP corpus contains 1064 utterances. Of these, 42 contained words that were not found in the CMU dictionary, so these were excluded. On 378 utterances, the number of syllables in the utterance differed between the STP annotation and the dictionary lookup; these were excluded as well. (Some such differences are to be expected, as there is sometimes variation in the number of syllables in a word; as an extreme example, *actually* is marked in the CMU dictionary as having four syllables but was in one case annotated as having only two.) This left a corpus of 644 utterances, 4735 words and 5916 syllables.

The stress-annotated portion of the BU corpus contains 114 sentences.[4] Sixteen of these were excluded because the number of syllables in the dictionary lookup and the annotations did not agree. Ten remaining sentences contained words that were not the in CMU dictionary; these words were added to the dictionary. Thus the final BU corpus contained 98 sentences, 1798 words and 2850 syllables.

Measuring Stress Regularity

Given a pattern of stresses, our method for calculating its regularity is simply to measure the variance in inter-stress intervals (ISIs). ISIs are measured by the distance in syllables between each pair of stresses. Thus two adjacent stressed syllables have a ISI of 1, two stresses separated by one unstressed syllable have an ISI of 2, etc. (One might prefer to define the distance between adjacent stressed syllables as 0. But this would not affect the results; it would simply subtract 1 from each ISI, and the resulting variance would be the same.) We measure ISIs only within each utterance in the STP corpus, and within each sentence in the BU corpus. (For simplicity, we will henceforth speak of "sentences" in both corpora.) Unstressed syllables before the first stress of a sentence and after the last one are simply ignored; here we follow the practice of earlier studies such as Roach (1982).

A difficult issue that we encounter here is the hierarchical nature of stress. In English, by most accounts, lexical stress patterns reflect at least three levels of stress; one or more additional levels of phrasal stress may also be added (for an alternative view, see Ladefoged 1993). Hayes (1995) argues for at least four levels of stress in English. For simplicity, consistency with the corpora being used, and consistency with prior studies of stress regularity (Roach 1982; Dauer 1983), we adopt a binary distinction here, marking completely unstressed syllables as "unstressed" (0) and marking all others as "stressed" (1). Hayes (1995) diagnoses a complete lack of stress by certain segmental properties such as the allowance of flapping on the preceding consonant. By this measure, unstressed vowels include schwa as well as some other vowels in certain contexts. This criterion seems to fit well with the labeling schemes of the BU and STP corpora and the CMU dictionary (assuming, in the case of the STP corpus and the CMU dictionary, that level 0 is unstressed and all other levels are stressed). For example, the suffix *-ing* in progressive participles is unstressed by Hayes's criteria, and is generally marked as such in the corpora and the dictionary; the final syllable of words like *elevate* or *participate* is considered stressed by Hayes' criteria and is (usually) marked as stressed in the data used here.

[4] The BU corpus data is presented in several different formats; of interest here is the ".aln" format, containing hand-corrected phonetic symbols and stress markers. The .aln files contain a corpus of 114 sentences, which was read (in whole or in part) three times, twice by one speaker (f2b) and once by a second speaker (f1a). The first reading by speaker f2b is the only one of the three readings that contains all 114 sentences, so this reading was chosen for use here.

A The first 10 valid utterances of the STP corpus

```
1. with what's going on today and i i think i think that we may not
   0    1      1 0   1  0 1   0   0 1 1     1 1     0    1  1   1
   0    1      1 0   0  0 1   0   0 0 1     0 1     0    0  0   1

2. okay well
   11   1
   11   1

3. yeah just because they're grandparents that doesn't automatically make them a
   0    1    0 0     1    0   1    1 0     0    1   0   1 0 1 0 0     1    0    0
   1    1    0 1     0        1    1 0     0    1   0   1 0 1 0 0     1    0    0

   good child carer
   1    1     1
   1    1     1 0

4. very very cognizant of and aware of all these type of
   1 0  1 1  1  0 0    0  0   01    0  1   0     1    0
   1 0  1 0  1  0 0    0  0   01    0  1   0     1    0

5. yeah they're like black corduroy
   1    0       1    1     1  0 1
   1    0       0    1     1  0 1

6. i put a stop to some of them as far as the door to door either religious groups
   1 1   0 1    0  1    1  0    0  1   0  0   1    0  1    1 0    1 1 0     1
   0 1   0 1    0  0    0  0    0  1   0  0   1    0  1    1 0    0 1 0     1

   or people
   0  1  1
   0  1  0

7. little prejudice course
   1 0    1  0 0    1
   1 0    1  0 0    1

8. she let me know in no uncertain terms that she wanted to use the bathroom
   1   1   1  1    0  1  0 1  0    1     0    1   1  0   0  1   0   1   1
   0   1   0  1    0  0  0 1  0    1     0    0   1  0   0  1   0   1   1

9. so on next night i spread the newspaper in the bathroom and she used them there
   1  1  1    1     1 1      0   1   1 0   0  0   1    1    0   1   1    0    1
   0  0  1    1     0 1      0   1   1 0   0  0   1    1    0   0   1    0    1

10.well right they destroyed it
   1    1     0    0 1       1
   1    1     0    0 1       0
```

B The first five valid sentences of the BU corpus

```
1. wanted chief justice of the massachusetts supreme court
   1  0   1     1  0    1  0   1   01  0     0 1     1
   1  0   1     1  0    0  0   1   01  0     0 1     1

2. in april the s j c's current leader edward hennessy reaches the mandatory
   0  10    0   1 1 1   1 0     1   0  1 0    1 0  0   1    0  0   1  0 1 0
   0  10    0   1 1 1   1 0     1   0  1 0    1 0  0   1    0  0   1  0 1 0

   retirement age of seventy and a successor is expected to be named in march
   0 1 0 0    1   0  1 0  0  0   0 0 1  0    0  0 1  0   0  0  1     0  1
   0 1 0 0    1   0  1 0  0  0   0 0 1  0    0  0 1  0   0  0  1     0  1

3. it may be the most important appointment governor michael dukakis makes during
   0  1   0  0   1    0 1  0    01     0    1 0  0   1 0     0 1 0   1     1 0
   0  0   0  0   1    1 1  0    01     0    1 0  0   1 0     0 1 0   1     1 0

   the remainder of his administration and one of the toughest
   0   0 1   0   0  1   0 1  01   0    0   1   0  0   1    0
   0   0 1   0   0  0   0 1  01   0    0   0   0  0   1    0

4. as w   b u r's margo melnicove reports hennessy will be a hard act to follow
   0  100 1 1 1   1  0  1  0 1    0 1     1 0  0   0    0  0 1    1   0  1 0
   0  100 1 1 1   1  0  1  0 1    0 1     1 0  0   0    0  0 1    1   0  1 0

5. in nineteen seventy six democratic governor michael dukakis fulfilled a campaign
   0  1   1    1 0  0  1   1 0 1  0   1 0  0   1 0     0 1 0   0  1      0 0  1
   0  1   1    1 0  0  1   1 0 1  0   1 0  0   1 0     0 1 0   0  1      0 0  1

   promise to depoliticize judicial appointments
   1  0    0  1 0 1 0 1    0 1 0    01     0
   1  0    0  1 0 1 0 1    0 1 0    01     0
```

Fig. 1 Samples of the two corpora used in the study, showing the text, the annotated stress patterns (above), and the dictionary stress patterns (below)

Table 1 Basic statistics on the STP corpus and BU corpus

	Sentences	Words	Syllables	% of syllables marked as stressed		% Agreement between annotations and citation forms
				Annotations	Citation forms	
STP corpus	644	4735	5916	54.0	44.5	75.5
BU corpus	98	1798	2850	48.8	44.3	88.9

Figure 1 shows a small sample of the data. The figure shows the first ten valid utterances of the STP corpus and the first five valid sentences of the BU corpus ("valid" meaning that they were eligible for use in our study). Below the text are shown, first, the annotated stress patterns, and second, the stress patterns taken from the CMU dictionary. All stress symbols have been relabeled as 1 or 0, as discussed above.

As a first step, it is instructive to compare the density of stresses in the annotations and the dictionary lookup (see Table 1). In both corpora, roughly half of all syllables are marked as stressed in both the annotations and the citation forms, though the annotations yield slightly more stressed syllables. We can also observe the degree of agreement between the annotations and the dictionary (see Table 1). On the STP corpus, the dictionary lookup and the annotations yielded matching stress symbols on 75.5% of syllables; on the BU corpus, the lookup and the annotations match on 88.9% of syllables. These measures give a general idea of how well the citation-form stress patterns of words predict their actual stress levels in natural speaking. The level of agreement seems to be considerably higher in formal English than in conversational English; the fairly low level of agreement on the STP corpus suggests that in spontaneous conversation, contextual stress adjustments occur quite frequently. One might wonder if a different classification of stress symbols into stressed and unstressed could improve the agreement between the annotations and the dictionary. On the STP corpus, we tried classifying annotation symbols of 0.5 and 0.25 as unstressed; this yielded agreement on 73.3% of syllables between the annotations and the dictionary lookup, slightly worse than the level of agreement reported above. (Only about 1% of syllables were labeled 0.25, so the categorization of these would make little difference.)

One possible problem with this measure of agreement is that the stresses may not always have been properly aligned. As noted earlier, cases in which the number of syllables for a sentence differed between the citation forms and the annotations were excluded, but there still may have been occasional alignment problems; if the annotations added a syllable at one point (in relation to the citation forms) and then deleted a syllable later, the overall syllable count would be the same and the sentence would be retained, but part of the sentence would be misaligned. Figure 1 contains an example of this; in the third sentence of the STP corpus, the annotations and the dictionary lookup yield the same number of syllables, but there are two differences in syllable tokenization which cancel out. The first 100 valid sentences in the STP corpus were inspected individually and only 3 sentences in this set exhibited this problem. Thus it seems to have been fairly uncommon, and probably accounts for only a small portion of the stress mismatches.

We now turn to the measurement of stress regularity. The first step is to randomize the stress pattern in each sentence, producing a random series of 0's and 1's with the same length and proportion of stressed syllables as the original sentence. We then measure the variance in inter-stress intervals on these random patterns. (In this test and in all tests below involving random choices, the test was run five times and the average value of the five runs was taken.) This gives a baseline measure of how much ISI variance there would be if there were absolutely no preference for regularity. Here we use the annotated stress values rather than the dictionary ones. (This matters only because the densities of annotated and dictionary stresses are not quite the same; we return to this point below.) We treat each ISI as a separate observation, thus ensuring that sentences with more syllables carry greater weight. For the STP corpus and the BU corpus, the random syllable patterns yield variances of 1.11 and 1.85, respectively; see the top row of Tables 2 and 3. The fact that ISI variance is higher in the BU corpus is probably due, first, to the somewhat lower density of stresses in that corpus, and secondly, to the much greater number of syllables per sentence (29.1 in the BU corpus versus 9.2 in the STP corpus), which allows for some quite long inter-stress intervals to occur by chance.

Next we measure ISI variance in the original annotated data, by the same method described above. (See the fifth row of Tables 2 and 3). This yields 0.79 for the STP corpus and 0.94 for the BU corpus. Since our interest is in the degree of regularity in comparison to a random pattern, we divide these values by the ISI variance of a random pattern as reported above. This yields "adjusted" variances of 0.71 for the STP corpus and 0.51 for the BU corpus. We can test the difference between the annotated stress data and a random stress pattern for significance, by finding the deviation of each ISI from the mean under both conditions; using a

Table 2 ISI Variance in the STP corpus

	ISI variance	Adjusted ISI variance	% Reduction in adjusted variance (in relation to previous row)
Random syllable order, annotated stresses	1.11	1.00	–
Random syllable order, citation-form stresses	1.68	1.00	–
Random word order, citation-form stresses	1.38	0.82	17.7**
Original word order, citation-form stresses	1.07	0.64	18.3***
Original word order, annotated stresses	0.79	0.71	−6.8*
Total variance reduction			29.2***

* $p < .01$; ** $p < .001$; *** $p < .0001$

Table 3 ISI variance in the BU corpus

	ISI variance	Adjusted ISI variance	% Reduction in adjusted variance (in relation to previous row)
Random syllable order, annotated stresses	1.85	1.00	–
Random syllable order, citation-form stresses	2.33	1.00	–
Random word order, citation-form stresses	1.36	0.58	41.5*
Original word order, citation-form stresses	1.14	0.49	9.4
Original word order, annotated stresses	0.94	0.51	−1.7
Total variance reduction			49.2*

* $p < .0001$

one-way ANOVA, we then compare these deviations across conditions. (As before, we adjust the deviations in proportion to the mean deviation for the random pattern.) In both corpora, the difference is highly significant: in the STP corpus, $F(1, 5096) = 39.21$, $p < .0001$; in the BU corpus, $F(1, 2582) = 74.08$, $p < .0001$.

The figures just reported indicate the overall degree of distributional stress regularity in the corpora, and thus indicate the effect of lexical, interlexical, and contextual stress patterning in combination. The next question is, how much do each of these individual factors contribute to distributional stress regularity? We begin by considering the effect of lexical patterning alone. We first arrange the words of each sentence in a random order. We then look up each word in the phonetic dictionary and find the stress pattern of its citation form. (As mentioned earlier, the dictionary was modified to make monosyllabic function words unstressed rather than stressed.) We then calculate the ISI variance of the resulting stress patterns. As before, we can measure the regularity of such a pattern by comparing its ISI variance to that of a random stress pattern. There is a complication here, however. It seems clear that the variance in stress distances will increase as the distances themselves increase. (Consider a stress pattern, and now imagine the same pattern with all of the distances doubled. Intuitively, the regularity of the two patterns is the same; but the second pattern will actually have *four* times the variance of the first, since the variance is the average of the squared deviations from the mean.) The density of stresses is somewhat lower in the citation-form patterns than in the annotations (particularly in the STP corpus), thus we would expect the ISI variance to be somewhat higher even if the degree of regularity was the same. Thus we randomize the citation-form stress patterns, find the variance (see the second row of Tables 2 and 3), and divide the ISI variance of our random-word-order citation-form patterns by this. The result is shown in the third row of Tables 2 and 3. The adjusted variances (0.82 for the STP corpus, 0.58 for the BU corpus) are less than 1, suggesting that lexical patterning reduces stress regularity. Using the method of comparing deviations as described above, we find that this effect is significant in both corpora: on the STP corpus, $F(1, 3990) = 12.58$, $p < .001$; on the BU corpus, $F(1, 2328) = 45.17$, $p < .0001$.

We now wish to examine the effect of interlexical patterning. We do this by once again using the stress patterns from the phonetic dictionary, but this time retaining the original order of the words. Again, we calculate the ISI variance and adjust this in proportion to the random-syllable-order citation-form variance. This yields adjusted variances of 0.64 on the STP corpus and 0.49 on the BU corpus. The difference between the variances of the original-word-order pattern and the random-word-order pattern indicates the contribution of interlexical patterning to stress regularity (compare the third and fourth rows of Tables 2 and 3). This difference is significant in the STP corpus, $F(1, 3990) = 16.28$, $p < .0001$, but not in the BU corpus, $F(1, 2328) = 2.51$, $p = .11$.

Finally, we examine the effect of contextual stress patterning. We can do this by simply comparing the variances of the original-word-order citation-form patterns and the original-word-order annotated patterns (using the adjusted variances in each case); compare the fourth and fifth rows of Tables 2 and 3. Note that, by this measure, the citation-form stresses actually have *lower* ISI variance than the annotated stresses on both corpora, suggesting that contextual patterning does not contribute to stress regularity. This difference is significant in the STP corpus, $F(1, 4543) = 10.09$, $p < .01$, but not in the BU corpus, $F(1, 2455) = 0.77$, $p = .38$.

Tables 2 and 3 also show the percentage reduction in variance due to each factor. It can be seen that lexical stress patterning reduces variance by 17.7% in the STP corpus and 41.5% in the BU corpus. Interlexical patterning causes a further reduction of 18.3% in the STP corpus and 9.4% in the BU corpus. Contextual patterning causes slight increases in variance in both corpora: 6.8% in the STP corpus and 1.7% in the BU corpus. The three factors in combination reduce ISI variance by 29.2% in the STP corpus and 49.2% in the BU corpus.

Discussion

The tests presented above reveal the contributions of lexical patterning, interlexical patterning, and contextual patterning to distributional stress regularity in English. The three factors in combination reduce ISI variance quite substantially, by almost 1/3 in the STP corpus and almost half in the BU corpus. The contribution of each factor can be seen by the decrease in ISI variance that occurs when it is added. When lexical patterning alone is present, variance is significantly reduced in both corpora (in comparison to a random stress pattern). Adding interlexical patterning produces a further reduction, which is significant in the STP corpus. Adding contextual patterning as well produces no further reduction in variance; indeed, the variance increases slightly (and significantly, in the STP corpus). We will discuss each of these findings in turn.

The large effect of lexical patterning on stress regularity, especially in the BU corpus, confirms the well-established idea—advanced by Chomsky and Halle (1968) and many later authors—that word-internal stress clash tends to be avoided. Much of this effect may be due to exactly the kind of phenomena discussed in metrical phonology—for example, stress shifts in lexical stems when a suffix is added, as in *inféstínfestátion*. One might wonder why the effect of lexical patterning is so much larger in the BU corpus than in the STP corpus. No doubt this is largely because the BU corpus has a much higher proportion of polysyllabic words than the STP corpus (36.8% vs. 18.6%). In a completely monosyllabic corpus, lexical stress patterning would have no effect: randomly permuting the words would be the same as randomly permuting the syllables, thus the two cases would yield equal variance. While the STP corpus is not completely monosyllabic, polysyllabic words are relatively infrequent, so that regularity of stress within them contributes rather little to overall stress regularity. (By contrast, the effect of interlexical patterning is larger in the STP corpus than in the BU corpus. Again, this is probably due to differences in word length; reordering the words of a sentence only affects ISIs that cross word boundaries, and these account for a much smaller proportion of ISIs in the BU corpus than in the STP corpus.)

Two further findings of the current study are perhaps more surprising. One is the significant effect of interlexical patterning on stress regularity. The role of word choice in stress regularity has received relatively little attention in stress research. Given that roughly half of all syllables are stressed—as suggested by all of the data presented here—it would seem that stress regularity is maximized by a pattern of alternation between strong and weak syllables (a point that has been observed in many studies of stress). This implies an avoidance of stress clashes—adjacent stressed syllables—as well as an avoidance of successive unstressed syllables. The question then is what kinds of specific phenomena in English tend to favor such an alternating pattern. While it is not yet possible to answer this question definitively, a few possibilities suggest themselves.

1. In general, since content words are usually stressed, stress clashes can be minimized by avoiding situations where content words are adjacent—that is, separating them by function words. Adjacent content words do not necessarily cause stress clashes, but they will do so if they have stressed syllables adjacent to the word boundary. While content word adjacencies are certainly not difficult to find in English, it may be that grammatical

rules and patterns of syntactic choice tend to reduce their frequency; a specific case of this will be seen below.

2. One observation made by Bolinger (1965) is that adjectives beginning with the prefix a—such as *asleep* and *awake*—can generally only occur predicatively (*John is asleep*; **the asleep boy*). Such adjectives are always iambic; since most nouns have initial stress, using such adjectives attributively would generally cause stress clashes.
3. As mentioned earlier, Kelly and Bock (1988) found that verbs tend to be preceded by stressed syllables while nouns tend to be followed by them; coupled with the tendency for nouns to be trochaic and verbs iambic, this may well tend to favor stress regularity. Kelly and Bock's findings deserve further investigation; one wonders what specific features of English grammar or usage give rise to the difference in stress context between nouns and verbs. One possibility is that it is related to the distribution of content words and function words; if nouns are more often preceded by function words and verbs are more often followed by them, this might give rise to the pattern found by Kelly and Bock. However, it is not at all obvious that this is the case. Nouns are very often preceded by determiners and prepositions; but verbs are often preceded by auxiliaries and pronouns (especially in informal speech). Both nouns and verbs are often followed by prepositions, though verbs—unlike nouns—are also often followed by determiners and pronouns. The difference in stress context between nouns and verbs is an intriguing finding that deserves further attention.
4. An important and widely studied locus of syntactic choice is the use of the complementizer *that* with embedded clauses (ECs), as in (5a) below, and object relative clauses (RCs), as in (5b).

   (5a) She said (that) John was coming
   (5b) The dog (that) John bought was black

   The complementizer is optional in such cases, and it is natural to wonder if its use might be affected by stress regularity. In one study involving EC constructions, Lee and Gibbons (2007) found that participants were more likely to use *that* when the following noun began with a stressed syllable, such as *Lucy*, rather than an unstressed syllable, such as *Louise*. It can be seen that this pattern promotes stress regularity. The stress pattern of the word sequence *said (that) Lucy* is most regular when the complementizer is included; the pattern of *said (that) Louise* is most regular when the complementizer is left out.
5. A further point with regard to ECs and object RCs concerns the distinction between full-NP and pronoun subjects:

   (6a) She said (that) she was coming
   (6b) She said (that) John was coming
   (6c) The dog (that) I bought was black
   (6d) The dog (that) John bought was black

   Corpus research has shown that *that* is much less likely to be included when the subject of the following clause is a pronoun, both in ECs and object RCs (Elsness 1984; Temperley 2003; Jaeger and Wasow 2005). Let us assume that the word preceding the dependent clause is generally a content word—a noun in the case of object RC constructions, a verb in the case of EC constructions. If the EC/RC subject is a full noun phrase, it may begin with a common or proper noun, as in (6b) and (6d) above, creating a content-word adjacency (though of course it may also begin with a determiner). Including *that* in such situations ensures that no content-word adjacency will occur. Conversely, if the EC/RC subject is a pronoun, as in (6a) or (6c), including *that* would guarantee two function words in a row, reducing stress regularity in another way.

It seems, then, that the more frequent use of *that* with full-NP dependent-clause subjects might enhance stress regularity. Note that we are not implying that this pattern in the use of *that* is due to the preference for stress regularity. Corpus and experimental studies have suggested, rather, that it is due mainly to discourse accessibility effects; *that* is more likely to be included when the referent of the dependent-clause subject is less available in the discourse—for example, a "new" entity that has not been previously mentioned (Ferreira and Dell 2000; Jaeger and Wasow 2005). Still, whatever its cause, it seems likely that this pattern has the effect of enhancing distributional stress regularity.

6. Another aspect of syntactic choice that may contribute to stress regularity is the ordering of conjuncts within a coordinate phrase. In a study by McDonald et al. (1993, Experiment 6), participants were read coordinate phrases of the form *[noun] and [noun]*: in each case, one of the nouns was disyllabic and the other was monosyllabic. In recalling the phrases, participants showed a slight tendency to put the disyllabic noun first when it was iambic, but not when it was trochaic. Assuming that *and* is normally unstressed, it can be seen that this pattern promotes stress regularity:

(7a)	iambic noun:	U S | and | S (regular)	S | and | U S (irregular)
(7b)	trochaic noun:	S U | and | S (irregular)	S | and | S U (regular)

The effect of stress pattern in McDonald et al.'s study was small, and appeared only when both nouns were inanimate; otherwise there was a strong preference to put the animate noun first and no effect of stress pattern. Still, this study suggests that stress regularity may have at least a small effect on the ordering of conjuncts.

The phenomena discussed above suggest that the effect of interlexical patterning on stress regularity may involve a large number of different phenomena of grammar and syntactic choice. Clearly, this area deserves further investigation.

The final finding of the current study is that contextual stress patterning does not increase stress regularity, and indeed slightly decreases it in conversational speech. This may seem surprising, in view of the considerable attention that has been given to contextual stress-clash avoidance in metrical phonology, especially in the work of Hayes (1984, 1995), We should note, however, that the current findings may not have much bearing on the kinds of stress-clash phenomena dicussed in metrical phonology. These phenomena mostly relate to distinctions among stressed syllables (Hayes 1995, pp. 19–20). According to Hayes, in a phrase like *thirteen men*, both syllables in *thirteen* are stressed; the effect of stress-class avoidance is to shift the main stress from the second stressed syllable to the first. Shifts of this kind might well not have been detected in the current study, since they concern higher-level stress distinctions. Thus, the absence of a contextual regularity effect observed in the current study cannot rule out the possibility that shifts of the kind discussed by Hayes occur in speech, perhaps quite frequently.

One way in which contextual stress patterning may favor regularity was mentioned earlier: The tendency to stress the second of three normally unstressed syllables. This tendency was examined in the corpora used here. The citation-form stress patterns were searched for patterns of exactly three unstressed syllables flanked by stresses on either side. For each such occurrence, the corresponding annotated patterns were found. Of interest were cases where just one of the three unstressed syllables was stressed in the annotations; the prediction was that the second of the three would be stressed more often than the first or third. The data is shown in Table 4; in both the BU corpus and the STP corpus, the predicted pattern is found.

Table 4 Annotations of the middle three syllables of the citation-form stress pattern "1 0 0 0 1"

Pattern	Count
STP corpus	
1 0 0	16
0 1 0	29
0 0 1	9
BU corpus	
1 0 0	7
0 1 0	14
0 0 1	7

A chi-square test (against an expected uniform distribution across the three patterns) showed that this was significant in the case of the STP corpus, $\chi^2 = 11.44$, $p < .01$, though not in the BU corpus, $\chi^2 = 3.5$, $p = .17$.

Of course, the fact that the ISI variance of the annotated stress patterns was higher than that of the citation-form patterns suggests that the deviations from citation-form stress patterns observed here do not, for the most part, represent strategies to increase regularity. The fact remains, however, that such deviations appear to be very common, especially in conversational speech; in the STP corpus, the annotated and citation-form stress values differed on almost 25% of all syllables. This raises the question, in what situations do such deviations arise, and how can they be explained? As noted earlier, some of these deviations may be due to imperfections in the labeling of citation-form stress patterns: for example, incorrect labeling of verb-noun homographs such as *pérmit*/*permít*. But inspection of the citation-form stresses suggests that only a very small percentage of syllables, probably less than 1%, were assigned incorrect citation forms (see the Appendix). One well-known kind of deviation from normal lexical stress patterns is that function words may be stressed under certain circumstances—for example, when in narrow focus or at the end of a phrase (Selkirk 1996). Indeed, an examination of the data shows that a majority of the deviations involved function words marked as unstressed in the dictionary but stressed in the annotations (57.6% of the deviations in the STP corpus were of this kind, 59.4% in the BU corpus). However, it appears from informal inspection of the data that very few of these stressed function words are either in narrow focus or in phrase-final position. These percentages also show that a substantial number of deviant syllables occur in content words—perhaps monosyllabic content words that are annotated as unstressed, or polysyllabic words whose annotated stress pattern differs from their citation form. Inspection of these deviations yields few obvious answers as to where and why they occur. One possibility, suggested by Cooper and Eady (1986), is that the conventional citation-form stress patterns for words do not always describe their normative use very well; Cooper and Eady report that *thirteen*—normally assumed to be iambic except in cases of stress-shift—is much more often pronounced as trochaic even in situations where stress-shift is not predicted (though, as noted earlier, it is really higher-level stress distinctions that are at issue here). Perhaps a significantly better match to spoken stress patterns could be achieved simply by using different citation forms. In any case, the very frequent deviations from citation-form stress patterns in conversational speech, both in function words and content words, deserve further study.

One (perhaps rather radical) interpretation of the data presented here is that stress regularity is to a large extent an optional, one might say decorative, aspect of speech which enhances its aesthetic effect but is inessential for communication. This would explain why it is found to a much greater degree in planned, formal speech (e.g. the BU corpus) than in spontaneous conversation (the STP corpus). The extreme case of this is of course poetry—a kind of language use which features a very high degree of stress regularity, and whose function is much more aesthetic than practical. In conversational speech, the factors influencing contextual stress patterns, whatever they may be, seem to carry much more weight than the preference for regularity—to the extent that they actually negate some of the regularity that would arise from normal lexical stress patterns. A comment by Hayes suggests that he might endorse this view of stress regularity as an inessential, largely "aesthetic" phenomenon:

> I suggest that the process of planning speech on-line does not evaluate the relative eurhythm [stress regularity] of competing potential outputs when “deciding” whether to apply rhythmic adjustment rules, but that given time to reflect, speakers can judge the rhythmic well-formedness of the result. (1995, p. 373)

To examine stress regularity in other languages would certainly be of interest. The results would probably vary quite considerably, due to the very different character of stress in different languages (Hayes 1995). Some languages have quite consistent placement of lexical stress—for example, on the initial syllable (as in Czech) or penultimate syllable (as in Polish). This might limit the potential for interlexical stress regularity; phenomena such as that discussed by Kelly and Bock (1988), in which the different stress environments of nouns and verbs are matched by different lexical stress patterns, would not be possible. In other languages, such as French, stress is primarily phrasal, with word-internal stress distinctions minimized or absent; the issue of stress regularity may simply not arise in such cases. In many other languages, however, word-internal stress distinctions are overt and variable, as in English; the issue of distributional stress regularity in such languages would be well worth exploring. Theoretical studies have observed stress-clash avoidance in a number of languages besides English (Nespor and Vogel 1989; Hayes 1995), but whether such phenomena contribute significantly to stress regularity is an open question. Of interest in this regard is a study by Wagner (2000), which presents an algorithm for predicting stress in German. Wagner found that stress could be predicted well using lexical, syntactic, and semantic information; she found no evidence of contextual stress shift.

A final issue that deserves consideration here is the role of the perceiver in stress. Certainly perceptual judgments of stress are strongly influenced and constrained by prosodic features such as F0, length, and intensity; this is clear from experimental work on stress perception (Fry 1958; Morton and Jassem 1965). But they may also be affected by the perceiver’s intuitions about normative stress patterns.[5] These intuitions might be affected by the same factors

[5] This point might be raised with regard to the stress annotations reported here: that is, they may have been affected by transcribers’ intuitions about normative stress patterns. However, this seems rather unlikely. The annotations were clearly done in a very careful and laborious manner, and in the case of the STP corpus, annotators were specifically instructed not to use knowledge of canonical stress patterns (Ostendorf et al. 1995; Hitchcock and Greenberg 2001). With regard to regularity, the fact that the stress annotations showed higher ISI variance than the citation-form stress patterns seems to rule out the possibility that the transcribers were imposing regularity on them to any significant degree. Another way of investigating this would be to examine multiple annotations of the same passage read by different speakers; if annotators were simply imposing their preconceived stress patterns, we would expect their annotations to differ very little across different speakers. As it happens, the BU corpus contains the same passage of text read by two different speakers (1963 syllables). The stress annotations for these passages reflect an agreement level of 91.1%; thus they differ quite considerably—almost as much as the annotated stress patterns in the BU corpus differ from the citation forms. This seems to make it clear that the annotators were not merely imposing preconceived stress patterns, and that they were highly influenced by the particular “performance” of the text that they were transcribing.
With regard to the role of the perceiver in stress regularity, an interesting study was done by Martin (1970). In this study, participants heard a sentence with an inserted pause of varying length, and had to judge the

that can affect stress patterns in production: in particular, knowledge of citation-form patterns and also of the conventions for deviating from citation form. (One wonders, for example, if *permit* would be more likely to be perceived as trochaic when used a noun and iambic when used as a verb.) Regularity may be a factor here as well. In this connection, consider once again the case of three unstressed syllables in a row:

(8) I gave it to the teacher.

Given the tendency towards regular stress patterning in English, perhaps listeners would tend to "hear" a stress on the second unstressed syllable (*to*), even if none of the three syllables was phonetically more prominent than the others. Thus, stress regularity in language may be a phenomenon of perception as well as production. This is an important point to bear in mind as we seek to gain a better understanding of the factors shaping distributional stress regularity and linguistic stress in general.

Appendix: Function and Content Words

The distinction between function (functional) words and content (contentive) words is well-established in theoretical and experimental linguistic work, and has proven important for studies of syntax, prosody, and language processing (see Cann 2001, for a survey). Several diagnostics for this distinction have been proposed. Content words are "open-class," admitting of new members, while function words are "closed-class"; new nouns frequently enter the language, while new determiners do not. Content and function classes also have different syntactic properties; content words assign thematic ("theta") roles while function words do not (Radford 1997). Content words invariably contain at least one stressed syllable, while function words may be completely unstressed (Selkirk 1996). Lexical and functional categories also may also differ in the way they are processed in comprehension, though this is controversial (Haveman 1996). These diagnostics prove to be largely convergent, though not entirely so (Cann 2001).

Our aim here is to identify precisely which English words are function words from a prosodic viewpoint—that is, words that are normally (or at least sometimes) unstressed. Here we encounter several problems.

1. Treatments of the function/content distinction do not always agree as to which categories of words are functional. There seems to be general agreement that nouns, verbs, and adjectives are contentive and that determiners, conjunctions, pronouns, and auxiliaries are functional. Prepositions are problematic: By syntactic criteria, they are sometimes considered content words (Radford 1997), but prosodically they are clearly function words as they are often unstressed. (This applies only to monosyllabic prepositions; polysyllabic prepositions like *over* and *above*, and indeed all polysyllabic words, do require stress [Selkirk 1996]). An even more problematic case is adverbs, which are often not mentioned at all in discussions of the function/content distinction. Manner adverbs are clearly open-class and in any case are almost always polysyllabic; other kinds of adverbs seem to be closed-class and in a few cases may be unstressed, such as the intensifier *so* and the sentential adverb *just*. Here we classify all adverbs as content words.

---

stress of syllables following the pause; perceived stress depended on the temporal position of syllables, with a
Footnote 5 continued
preference for positions that maintained an even spacing of stresses with the prior context. This might be taken as evidence that the preference for stress regularity can affect the perception of stress. In Martin's experiment, however, the context sentences were read with perfectly isochronous stresses (following an isochronous click track), thus it was a rather artificial situation.

2. In some cases, there is even uncertainty as to the syntactic category of a word. For example, words like *here* and *there* lack objects and thus are like adverbs, but they can modify nouns (*the restaurants here are amazing*) and thus are like prepositions; sometimes they are regarded as "pro-adverbs." The word *near* takes an object, like a preposition, but can also take an intensifier (e.g., *very*), like an adjective. Such words were classified based on what seems to be their normal stress pattern; *here* and *there* are normally stressed, while *near* may be unstressed.
3. English features a number of homographic word pairs with differing stress patterns; in most such cases the two words are in different syntactic categories. (The CMU dictionary sometimes has multiple pronunciations for words but does not have them labeled by syntactic category, so this is of no help.) These include noun/verb homographs such as *permit*; words like *on* which are unstressed as prepositions but stressed as particles; forms of *have*, which are stressed as main verbs but unstressed as auxiliaries; and *that*, which is unstressed as a complementizer but stressed as a demonstrative. Also problematic is *wh*-words, which are stressed when used as interrogative markers but sometimes unstressed in other contexts (*The man who left*; *I laughed when I saw him*). In such cases, words were simply labeled according to what seemed to be their most common usage.

With these considerations in mind, the following list of function words was constructed.

Prepositions and Conjunctions: and as at but by down from for if in like near nor of off on or out per plus since so than through to up while with worth
Pronouns: he her him his I it its me my one she their them they us we you your
Determiners: a an each no some that the these this those
Auxiliaries and Modals: am are be been is was were can can't could may might must should will won't would
Subject-Verb Contractions: he'd he'll he's I'd I'll I'm it's I've she'd she'll she's they'll they're they'd they've we'd we'll we're we've you'd you'll you're you've

Some might disagree with certain choices; and again, this purely orthographic system is bound to make some errors in cases of differently-stressed homographs. The question is how much these problems affect the labeling of word tokens. It seems that the vast majority of word tokens are unambiguous and uncontroversial. The reader is invited to consider the samples of the two corpora in Fig. 1 to see if any of the citation-form stress patterns listed there seem incorrect. We would argue that only one syllable (out of 302 syllables in the two samples) is given an incorrect citation form: The word *on* in sentence 1 of the STP corpus is a particle, not a preposition, and should therefore be stressed. Thus it seems unlikely that such mislabeled or controversial cases would have any significant effect on the results reported here.

References

Abercrombie, D. (1967). *Elements of general phonetics*. Chicago: Aldine.
Beckman, M. E., Swora, M. G., Rauschenberg, J., & de Jong, K. (1990). Stress shift, stress clash, and polysyllabic shortening in a prosodically annotated discourse. In *Proceedings of the First International Conference on Spoken Language Processing* (pp. 5–8).
Bolinger, D. (1965). *Forms of English: Accent, morpheme, order*. Cambridge: Harvard University Press.
Cann, R. (2001). Functional versus lexical: A cognitive dichotomy. In R. D. Borsley (Ed.), *The nature and function of syntactic categories* (pp. 37–78). New York: Academic Press.
Chomsky, N., & Halle, M. (1968). *The sound pattern of English*. New York: Harper and Row.
Cooper, W. E., & Eady, S. J. (1986). Metrical phonology in speech production. *Journal of Memory and Language, 25*, 369–384.

Cruttenden, A. (1997). *Intonation*. Cambridge, UK: Cambridge University Press.
Dauer, R. M. (1983). Stress-timing and syllable-timing reanalyzed. *Journal of Phonetics, 11*, 51–62.
Elsness, J. (1984). *That* or zero? A look at the choice of object clause connective in a corpus of American English. *English Studies, 65*, 519–533.
Ferreira, V. S., & Dell, G. S. (2000). Effect of ambiguity and lexical availability on syntactic and lexical production. *Cognitive Psychology, 40*, 296–340.
Fry, D. B. (1958). Experiments in the perception of stress. *Language and Speech, 1*, 126–152.
Greenberg, S. (1997). The Switchboard Transcription Project. *1996 Large Vocabulary Continuous Speech Recognition Summer Research Workshop Technical Report Series*, Research Report #24. Baltimore, MD: Center for Language and Speech Processing, Johns Hopkins University.
Haveman, A. P. (1996). Effects of frequency on the auditory perception of open- versus closed-class words. In *Proceedings of the Fourth International Conference on Spoken Language Processing* (pp. 78–81).
Hayes, B. (1984). The phonology of rhythm in English. *Linguistic Inquiry, 15*, 33–74.
Hayes, B. (1995). *Metrical stress theory*. Chicago: University of Chicago Press.
Hitchcock, L., & Greenberg, S. (2001). Vowel height is intimately associated with stress accent in spontaneous American English discourse. In *EUROSPEECH-2001* (pp. 79–82).
Jaeger, T. F., & Wasow, T. (2005). Processing as a source of accessibility effects on variation. In *Proceedings of the 31st Annual Meeting of the Berkeley Linguistics Society*.
Kelly, M. H., & Bock, J. K. (1988). Stress in time. *Journal of Experimental Psychology: Human Perception and Performance, 14*, 389–403.
Ladefoged, P. (1993). *A course in phonetics*. Fort Worth: Harcourt Brace.
Lee, M., & Gibbons, J. (2007). Rhythmic alternation and the optional complementiser in English: New evidence of phonological influence on grammatical encoding. *Cognition, 105*, 446–456.
Lehiste, I. (1972). The timing of utterances and linguistic boundaries. *Journal of the Acoustical Society of America, 51*, 2018–2024.
Liberman, M., & Prince, A. (1977). On stress and linguistic rhythm. *Linguistic Inquiry, 8*, 249–336.
Martin, J. G. (1970). Rhythm-induced judgments of word stress in sentences. *Journal of Verbal Learning and Verbal Behavior, 9*, 627–633.
McDonald, J. L., Bock, K., & Kelly, M. H. (1993). Word and world order: Semantic, phonological, and metrical determinants of serial position. *Cognitive Psychology, 25*, 188–230.
Morton, J., & Jassem, W. (1965). Acoustical correlates of stress. *Language and Speech, 8*, 159–181.
Nespor, M., & Vogel, I. (1989). On clashes and lapses. *Phonology, 6*, 69–116.
Ostendorf, M., Price, P. J., & Shattuck-Hufnagel, S. (1995). *The Boston University Radio News Corpus*. Boston University Technical Report No. ECS-95-001.
Pike, K. (1945). *The intonation of American English*. Ann Arbor: University of Michigan Press.
Radford, A. (1997). *Syntactic theory and the structure of English*. Cambridge, UK: Cambridge University Press.
Roach, P. (1982). On the distinction between ‘stress-timed’ and ‘syllable-timed’ languages. In David Crystal (Ed.), *Linguistic controversies* (pp. 73–79). London: Edward Arnold.
Selkirk, E. (1984). *Phonology and syntax: The relation between sound and structure*. Cambridge: MIT Press.
Selkirk, E. (1996). The prosodic structure of function words. In J. L. Morgan & K. Demuth (Eds.), *Signal to syntax: Bootstrapping from speech to grammar in early acquisition* (pp. 187–213). Mahwah, NJ: Lawrence Erlbaum.
Sluijter, A. M. C., & van Heuven, V. J. (1996). Spectral balance as an acoustic correlate of linguistic stress. *Journal of the Acoustical Society of America, 100*, 2471–2485.
Temperley, D. (2003). Ambiguity avoidance in English relative clauses. *Language, 79*, 464–484.
Wagner, P. S. (2000). Evaluating metrical phonology—a computational-empirical approach. In *Proceedings of KONVENS (Konferenz zur Verarbeitung natürlicher Sprache)* (pp. 243–248).